Motivation and Problem Definition:

Social Network websites becomes the major environment for the users to communicate and express their opinions. But, there are some users use the social network websites in negative way like cyberbullying.

The united states government defines the Cyberbullying as a bullying that takes place over digital devices like cell phones, computers, and tablets. And, it can occur through SMS, Text, and apps, or online in social media, forums, or gaming where people can view, participate in, or share content. Cyberbullying includes sending, posting, or sharing negative, harmful, false, or mean content about someone else. It can include sharing personal or private information about someone else causing embarrassment or humiliation [1]. Moreover, multiple studies mention that cyberbullying can affect any anyone especially young people and they are at high risk. For example, National Center for Education Statistics and Bureau of Justice Statistics reported that 9% of students in grades 6–12 experienced cyberbullying [2]. Moreover, Kosciw, J. G. et al. say that 55.2% of LGBTQ students experienced cyberbullying [3].

Cyberbullying has negative consequences on victim. For example, Sourander et al. say that cyberbullying leads to serious pathological experience such as depression, self-harm and suicide attempt [4]. In addition, S. Hinduja and J. W. Patchin say that the effects of cyberbullying can start from temporary anxiety to suicide [5]. Kowalski and limber mention that 90% of the young people victims don't tell their parent about their cyberbullying experience [6].

Manual detection of cyberbullying is very difficult because the huge volume of data on social network websites. Therefore, accurate and automated detection of cyberbullying is more effective. Several studies focused mainly on the content-based feature such as profanity, pronouns, bags of word, term frequency inverse document (TFIDF) and cyberbullying words. For example, Foong and Oussalah present an automated cyberbullying system detection. The system is based on natural language processing, text mining and machine learning to detect cyberbullying. The authors employed different textual features for the classifier such TF-Idf, linguistic Inquiry, word count features and Dependency features. In addition, the authors conduct their experiment on the collected dataset from ASKfm website [7]. Little studies are focused on user-based feature. Therefore, in this project, we will focus on the user-based feature and we will try to extract new user-based feature to enhance the accuracy of detection cyberbullying.

Methodology:

Dataset Description

In this project, we will use the dataset on cyberbullying which is collected by impermuim on Kaggle (https://www.kaggle.com/c/detecting-insults-in-social-commentary/data). The dataset is included six files. we will use the dataset in train.csv file. the total number of samples in the dataset is 3947. We will divide the dataset into 3000 samples train dataset and 947 samples test dataset.

Dataset Preprocessing

The dataset preprocessing stage includes noise removal (removing HTML,XML and metadata), tokenization means that breaking the stream of text into small units called token. Normalization which includes removing stop words, converting the uppercase into lower case, non-ASCII characters and stemming.

Feature extraction

We will choose the textual features that will be applied for classifiers. Two types of feature extraction will be applied count vector feature and TF-IDF feature.

References:

[1] "What is Cyberbullying" Accessed March 15,2018. [online]. Available: https://www.stopbullying.gov/cyberbullying/what-is-it/index.html

[2] National Center for Education Statistics and Bureau of Justice *Statistics*, *2011*.

[3] Kosciw, J. G., Greytak, E. A., Bartkiewicz, M. J., Boesen, M. J., & Palmer, N. A, "The 2011 National School Climate Survey: The experiences of lesbian, gay, bisexual and transgender youth in our nation's schools*". New York: GLSEN.*

[4] Sourander, A., Brunstein-Klomek, A., Ikonen, M., Lindroos, J., Luntamo, T., Koskelainen, M., Ristkari, T., Hans Helenius, H. 2010. "Psychosocial risk factors associated with cyberbullying among adolescents: A population-based study." *Arch Gen Psychiatry*, 67: 720-728.

[5] S. Hinduja and J. W. Patchin. "Bullying, cyberbullying, and suicide". *Archives of suicide research*, 14(3):206–221, 2010.

[6] Robin M Kowalski and Susan P Limber." Electronic bullying among middle school students". *Journal of adolescent health*, 2007,41(6):S22–S30.

[7] Y. J. Foong and M. Oussalah, "Cyberbullying System Detection and Analysis," *2017 European Intelligence and Security Informatics Conference (EISIC)*, Athens, 2017, pp. 40-46.
doi: 10.1109/EISIC.2017.43